WHAT IS CLAIMED IS:

1.   A method for structuring video by probabilistic merging of video segments, said method comprising the steps of:

a) obtaining a plurality of frames of unstructured video;

b) generating video segments from the unstructured video by detecting shot boundaries based on color dissimilarity between consecutive frames;

c) extracting a feature set by processing pairs of segments for visual dissimilarity and their temporal relationship, thereby generating an inter-segment visual dissimilarity feature and an inter-segment temporal relationship feature; and

d) merging video segments with a merging criterion that applies a probabilistic analysis to the feature set, thereby generating a merging sequence representing the video structure.

2.   The method as claimed in claim 1 wherein step b) comprises the steps of:

generating color histograms from the consecutive frames;

generating a difference signal from the color histograms that represents the color dissimilarity between consecutive frames; and

thresholding the difference signal based on a mean dissimilarity determined over a plurality of frames, thereby producing a signal that indicates an existence of a shot boundary.

3.   The method as claimed in claim 2 wherein the difference signal is based on a mean dissimilarity determined over a plurality of frames centered on one of the consecutive frames and corresponding in number of frames to a fraction of the frame rate of video capture.

4.   The method as claimed in claim 2 further including the step of morphologically transforming the thresholded difference signal with a pair of

structuring elements that eliminate the presence of multiple adjacent shot boundaries.

5.   The method as claimed in claim 1 wherein the processing of pairs of segments for visual dissimilarity in step c) comprises the steps of computing a mean color histogram for each segment and computing a visual dissimilarity feature metric from the difference between mean color histograms for pairs of segments.

6.   The method as claimed in claim 1 wherein the processing of pairs of segments for their temporal relationship in step c) comprises the processing of pairs of segments for a temporal separation between pairs of segments and for an accumulated temporal duration between pairs of segments.

7.   The method as claimed in claim 1 wherein step d) comprises the steps of:

generating parametric mixture models to represent class-conditional densities of the inter-segment features comprising the feature set; and

applying the merging criterion to the parametric mixture models.

8.   The method as claimed in claim 7 wherein step d) is performed in a hierarchical queue and comprises the steps of:

initializing the queue by introducing each feature into the queue with a priority equal to the probability of merging each corresponding pair of segments;

depleting the queue by merging the segments if the merging criterion is met; and

updating the model of the merged segment and then updating the queue based upon the updated model.

9.   The method as claimed in claim 1 wherein representing the merging sequence is represented in a hierarchical tree structure.

10.   A computer storage medium having instructions stored therein for causing a computer to perform the method of claim 1.

11.   A method for structuring video by probabilistic merging of video segments, said method comprising the steps of:

a) obtaining a plurality of frames of unstructured video;

b) generating video segments from the unstructured video by detecting shot boundaries based on color dissimilarity between consecutive video frames;

c) extracting a feature set by processing pairs of segments for visual dissimilarity and their temporal relationship, thereby generating an inter-segment visual dissimilarity feature and an inter-segment temporal relationship feature;

d) generating a parametric mixture model of the inter-segment features comprising the feature set; and

e) merging video segments with a merging criterion that applies a probabilistic Bayesian analysis to the parametric mixture model, thereby generating a merging sequence representing the video structure.

12.   The method as claimed in claim 11 wherein the processing of pairs of segments for visual dissimilarity in step c) comprises the steps of computing a mean color histogram for each segment and computing a visual dissimilarity feature metric from the difference between mean color histograms for pairs of segments.

13.   The method as claimed in claim 11 wherein the processing of pairs of segments for their temporal relationship in step c) comprises the processing of pairs of segments for a temporal separation between pairs of segments and for an accumulated temporal duration between pairs of segments.

14.   The method as claimed in claim 11 wherein the parametric mixture model generated in step d) represents class-conditional densities of the inter-segment features comprising the feature set.

15. The method as claimed in claim 11 wherein step e) is performed in a hierarchical queue and comprises the steps of:

initializing the queue by introducing each feature into the queue with a priority equal to the probability of merging each corresponding pair of segments;

depleting the queue by merging the segments if the merging criterion is met; and

updating the model of the merged segment and then updating the queue based upon the updated model.

16. The method as claimed in claim 11 wherein the merging sequence is represented in a hierarchical tree structure that includes a frame extracted from each segment and displayed at each node of the tree.

17. A computer storage medium having instructions stored therein for causing a computer to perform the method of claim 11.

18. A method for structuring video by probabilistic merging of video segments, said method comprising the steps of:

a) obtaining a plurality of frames of unstructured video;

b) generating video segments from the unstructured video by detecting shot boundaries based on color dissimilarity between consecutive video frames;

c) extracting a feature set by processing pairs of segments for visual dissimilarity and their temporal relationship, thereby generating an inter-segment visual dissimilarity feature and an inter-segment temporal relationship feature;

d) merging adjacent video segments with a merging criterion that applies a probabilistic Bayesian analysis to parametric mixture models derived from the feature set, thereby generating a merging sequence; and

e) representing the merging sequence in a hierarchical tree structure.

19.   The method as claimed in claim 18 wherein representing the merging sequence in a hierarchical tree structure includes displaying a frame extracted from each segment.

20.   A computer storage medium having instructions stored therein for causing a computer to perform the method of claim 18.